

研究報告書

「中脳神経回路網による価値情報の形成機構の解明」

研究タイプ: 通常型

研究期間: 平成 21 年 10 月～平成 25 年 3 月

研究者: 小林 康

1. 研究のねらい

臨機応変に柔軟でかきこい制御、情報通信の技術基盤として、生物に模した「自律探索型の学習機械」の開発が必要とされている。生物らしい情報処理の基本原則である、個体個体の経験に基づく「内的動機付け(やるき)」、「未来予測(報酬、コストやリスクの見通し)」の自律探索的な学習機構を神経生理学的レベルからアルゴリズムレベルにわたって解明し、「生き物らしさ」を機械に組み込むことをねらいとした。「目は口ほどに物を言う」ということわざを裏付けるように、様々な精神・神経疾患によって眼球運動に障害が生じる。本研究では、この眼球運動を定量的に測定可能なバイオマーカーとして活用し、被験者が患う精神・神経疾患の病態を誰でも簡単に評価することを可能にする新しい診断法の開発を目指した。

2. 研究成果

(1) 概要

サルとヒトの神経生理学研究、ヒトの心理物理研究により、意志・行動決定の神経回路機構を解明した。

- a) 強化学習の神経機構: 生物の意志決定の基本原則は報酬獲得であり、その行動制御の根底となる報酬の予測学習(強化学習)の神経回路機構、計算メカニズムについて覚醒サルを用いた、神経生理学的研究によって明らかにした。
- b) 隠れた運動行動決定過程を目の動きから読み取る: 神経疾患によって生じるヒトの「こころの戸惑い、震え」を眼球運動から読み解くことを目指し、注意障害多動の児童、健常被験者を対象に眼球運動を計測するヒトの心理物理研究を行った。そして、視線移動に関わるサッケードの反応時間、注視中の固視微動の解析により、運動の実行、経験に基づく適切な随意運動発現に必要な運動準備状態(見た目にはわからない行動決定過程)が眼球運動より読み取れることがわかった。

(2) 詳細

a) 強化学習の神経機構:

強化学習は手がかり刺激に対して行動の報酬を予測し、予測と実際に得られた報酬との差、「報酬予測誤差」を最小化するように予測を随時更新して、最大報酬を得る行動を最適化することであり、学習アルゴリズムの理論研究のみならず、強化学習が実装されている生理学的ハードウェア研究が盛んに行われている。現在、中脳ドーパミンニューロン(DAcell)が報酬予測誤差を表現しているということがほぼ確立されているが、計算理論の重要な鍵となる誤差信号の計算メカニズムについてその実体が明らかにされていない。DAcellは大脳基底核などから抑制性、大脳皮質などから興奮性の入力を受けるが、中脳のアセチルコリン作動性の脚橋被蓋核(PPTN)がDAcellに対してもっとも強力な興奮性入力を送っている。このことから、PPTNがDAcellによる報酬予測誤差計算の中心であることが示唆される。我々はモンキーチェアに座ったサル眼球運動をサーチコイル法で計測し、複数の計算機を連動させ、課題制御、データ取り込みを行うシステムを構築し、ジュース報酬で眼球運動課題を訓練したサルに報酬予測サッケード課題を行わせ、PPTNからニューロン活動記録を行っている。その結果、1)サルに課題開始刺激を呈示すると活動が始まり、手がかり刺激による予測報酬量が多いと活動が大きくなるような報酬が与えられるまで持続するニューロン活動(参考文献 1,2,3) 自発レベルでは比較的高頻度で発火し、課題開始から活動が減少して多くの報酬量が予測される試行ではより活動が減少するという1)とはちょうど鏡像関係にあるニューロン活動(論文投稿中)、3)報酬予測とは無関係に実際にサルに報酬が与えられると活動の増加が起こり与えられた報酬量が多いとより活動を増加させるような一過性のニューロン活動(参考文献 1,3)がそれぞれ独立したニューロン群から得られた。以上の結果から計算モデルで予言された報酬予測誤差計算に必要な、「持続的な興奮性の記憶された予測報酬の情報」「持続的な抑制性の記憶された予測報酬の情報」と「短時間の興奮性の実際に得られた報酬の情報」というすべての要素が、それぞれ分離独立した形でサル中脳PPTNに表現され、報酬予測誤差が計算されているということが明らかになった。

これら報酬関連活動の他に、PPTNでサルのサッケード運動時や視覚刺激呈示に対して一過性の活動上昇がみられた。これらは、約半数のニューロンで課題の文脈に対する依存性があった。これらの活動は学習された行動を文脈に従って遂行するのに役立っているものと思われる(参考文献 2)。また、状況依存的に10Hz程度の低頻度で周期的に活動するPPTNのニューロン群(おそらくアセチルコリン作動性ニューロン、サルが課題を始めると活動が周期的になる)でサッケード運動に伴う活動停止がみられた。これらのニューロンではサッケードの方向・振幅・課題文脈によらず、サッケード運動の開始に先立って活動が停止し、その停止はサッケード終了後まで続いた。外側膝状体がPPTNのアセチルコリン作動性入力を受けることから、この活動はサッケード抑制(サッケード運動が起こる前に視覚情報を遮断し、予測的に視覚像のぶれを除去する)と関係していると思われる(論文投稿準備中)。このように、DAcellと関係の強い中脳PPTNは報酬予測誤差計算のみならず、状況依存的な行動の実行や、円滑な感覚情報取り込みに関係していると思われる。

外部との研究交流:嫌悪と選好の脳内相互関係のモデル化:

中脳で主に報酬、覚醒、持続的注意に関係する、アセチルコリン作動性核「脚橋被蓋核」(我々が記録している)と嫌悪、衝動性、時間割引に関係するセロトニン作動性ニューロン核「背側縫線核」(関西医大中村加枝教授(さきがけ脳情報の解読と制御研究員)が記録している)のサル眼球運動課題中に得られたニューロン活動について、異なる研究室で得られた実験データを同じ解析方法で解析することを試み同じフォーマットの図を作ったうえでお互いのデータの比較検討を行い、報酬、嫌悪と強化学習に対する役割を明らかにするモデルを提唱した(参考文献3)。それぞれのニューロン活動の違いより、事象を予測する「タイムスケール」が脚橋被蓋核システム(選好)と背側縫線核システム(嫌悪)とでは異なる(いやなことの方が持続する(背側縫線核システムの方が情報の時間スケールが長い))ことが明らかになった(参考文献)。また、手がかり刺激によってジュース報酬を予期させる眼球運動課題をサルに行わせたところ、脚橋被蓋核と背側縫線核において、正負両方の符号で課題報酬の期待に関係する時間経過の類似したニューロン活動がみつかった。これらの結果は、アセチルコリン、セロトニンといった、いわゆるニューロモジュレーターによって、嫌悪と報酬といった「相反する」情報が中脳レベルにおいて様々なスケールで一元化されていることを示唆する(参考文献4)。

b) 隠れた行動決定過程を目の動きから読み取る:さまざまな神経疾患によって生じるヒトの「こころの戸惑い、震え」を眼球運動から読み解くことを研究目的とした。こどもの注意障害多動、高齢者のパーキンソン病はともに中脳、大脳基底核障害によって起きることが多いが、障害はそれぞれ異なる行動の異常として現れ、病態の神経メカニズムは未解明である。本研究では神経回路がよく研究されている眼球運動制御のシステム動態を手がかりにして神経疾患の早期発見、病状の進行度の生理学的理解、病態解明を目指した。具体的には視線移動に関わるサッケードの反応時間(標的が呈示されてから眼球が動き始めるまでの時間)を行動遂行の指標にし、運動を準備している注視中(見た目にはわからない行動決定過程)の固視微動から、適切な運動制御に必要な運動準備状態を読み取ることを試みた。

以下の心理物理課題を被験者に行わせた。被験者の眼球運動は、被験者の眼前のコンピューターディスプレイ画面に呈示された光点を一定時間注視させ、次々と現れる光点に向かって視線を移動させることで誘発させた。なお、被験者の頭部は顎台で固定され、眼球運動は非侵襲、高速度の眼球運動測定装置(阪大病院に設置)により測定した。

結果1) 注意障害多動患者では、注視に異常が見られ(光点をじっと注視するのが困難)、視線移動に関わるサッケードの反応時間に健常被験者と比べて遅れがみられた。さらに、サッケードの反応時間を短くする効果のあるギャップ課題(注視中に注視点を消灯させ、その後標的を呈示させるので運動の準備が可能)にしてもなお反応時間が短縮されない(ギャップ効果はみられるものの非常に弱い)ことが観察された。眼球を安定化させる注視システムには随意的なもの(随意的に起こさなければならないタイプの注視の解放)と自動的なもの(ギャップによって自動的に起きる注視の解放)があることが知られているが、この結果は注意障害多動患者がサッケードを遂行する際はむしろ、随意的な注視の解放が十分にできずに反応時間が長いことが示唆された。

結果2)たとえば、パーキンソン病などにより、様々な随意的な眼球運動[例:随意的なサッカー(呈示された光点への反射的な視線移動を抑制して光点の反対位置に視線を移動させるアンチサッカーなどに異常が見られる(反応時間が遅くなる、アンチサッカーを失敗する)、また、注視中の固視微動にも異常が生じることが知られている。注視中の固視微動は視覚情報処理に必須であることが最近明らかになっているが、サッカー遂行前の注視中の固視微動、ターゲットへのアンチサッカーから適切な随意サッカーの実行に必要な運動準備状態を読み取った。

健常被験者にアンチサッカーと通常のサッカー(プロサッカー課題)課題を行わせ、サッカー、あるいは注視中の固視微動を記録し、サッカーの反応時間、課題の成功失敗、固視微動の頻度、大きさ、瞬きの頻度、瞳孔径などを運動準備の度合い、衝動性の指標として定量解析を行った。テストを行った結果、あらかじめ運動準備が必要なアンチサッカー課題では、サッカーの反応時間が遅くなることが確認されさらに、アンチサッカーの準備状態の注視中は、プロサッカーの時と比べて、固視微動が減少していることがわかった。また、被験者がアンチサッカーを失敗する(意識には上らない)ときには、準備中の固視微動が異常に増加していることがわかった。

固視微動を随意的に押さえることが、「あっち向けほい」などの反射抑制を適切に行うときに必要なことがわかった(参考文献5)。

以上、固視微動を解析することによりこれからおこる運動の結果を実時間で予測できることがわかった。現在、固視中の固視微動、瞬き、瞳孔サイズ(ともに視覚認知に必須であり、注意など精神活動と深い関係がある)とサッカー制御、衝動性、脳内セロトニン濃度、発達の効果の関係、脳疾患患者(パーキンソン病、本態性振戦、注意障害多動、進行性核上性麻痺)を被験者にして、データを解析しているところである。

3. 今後の展開

強化学習の神経生理学研究の将来展望:

生物の「報酬・嫌悪刺激の価値評価をし、行動選択をする」神経機構を明らかにしたい。そのために、眼球運動課題を行う霊長類を動物モデルとして、視覚刺激—運動学習課題における、脳幹のセロトニン系の背側縫線核、アセチルコリン系の脚橋被蓋核、および中脳ドパミン系という異なる神経伝達物質を制御する領域の神経活動の変化を計測する。各領域が予測的価値・行動・報酬価値信号をどの様に計算し表現しているのか、その並列分散的かつ統合プロセスによりどのように行動決定、予測更新に至るのかを神経生理学的に明らかにする。三つの神経伝達物質の相互作用による価値情報にもとづく行動決定学習モデルは、従来のドパミン系に限局した報酬学習に加え、嫌悪情報や行動の種類も含めた幅広い予測学習モデルを構築することに寄与する。我々は、慢性サルを用いた神経生理学的研究により、脳脚橋被蓋核(PPTN)における報酬予測誤差計算機構を明らかにしてきた(参考文献)。この研究を進めることで脳の長期記憶(シナプス記憶)の書き込み、読み出し、短期記憶(ワーキングメモリ; 大脳、脳幹を含む反響回路循環による神経情報の短期的保持)、ニューロンによる(学習誤差)信号計算機構といった脳科学、情報科学の重要な問題にメスをいれるとともに、報酬嫌悪の予測(行動の動機付け、危険の事前回避)と報酬を獲得するまでずっと報酬予測をし



続ける時間との関係(価値の時間的割引効果;いつまでも予測した報酬がこないと予測報酬価値が低下する、せっかちさ)の脳内機構を神経生理学的に解明し、価値、予測と時間の情報が脳内でどのように表現されているのかといった、学習アルゴリズムレベルの研究に到達したい。最終的にはこれらの神経生理学的研究を進展させ、マイクロからマクロレベルでの社会行動経済学の発展(どうして人間はこんなにせっかちなのか、また様々な国民性の違いを説明する社会環境、脳内機構の解明)に寄与が期待されるレベルまで研究を進展させていきたいと思う。

ヒト心理物理実験の将来展望:

眼球運動は制御対象が単純であり、厳密かつ定量的な計測に特別な技術を必要としない。神経生理学研究に基づく分析から、眼球運動が疾患のバイオマーカーとして有効であることが今後確立されれば、この簡便で安価な技術を広く普及させ、様々な神経疾患の症状の発現を予防する方法の開発や、各患者の神経疾患症状の進行に合わせた治療法の計画など様々な応用分野で貢献できればと思う。

4. 自己評価

臨機応変に柔軟でかきこい制御の学習機構を神経生理学的レベルからアルゴリズムレベルにわたって解明することを目指した。さきがけ研究により、強化学習に関連した脳幹ニューロンで正と負の2方向性の(内的)動機付け、(外的情報による)報酬予測の情報表現が得られた。さらに、報酬予測情報は内的動機付け情報の上に重畳されていることがわかった。この結果は、「自分でやる気を起こさなければ学習がきちんと進まない」ことや「自分で嫌気を起こさなければ、いやなことは忘れられない」ことを示唆すると思われる。

将来的に、内的に生成させる正と負のやる気、嫌気、といった内的予測要素に外界の外的環境要素を組み込んだ学習理論を構築し、自発的で効率よい情報通信の適応的制御システム構築に生かしていきたい。

また、眼球運動を定量的に測定可能なバイオマーカーとして活用し、被験者が患う精神・神経疾患の病態を誰でも簡単に評価することを可能にする新しい診断法を開発を目指したが、さきがけ研究による「固視微動の計測」により、当初予測もしていなかった、「視線の非常に小さな揺らぎ」によって、オンラインで運動準備状態や脳の状態をモニターできる可能性を得ることができた。今後はこのような精密な眼球運動や各種計測により脳疾患の早期発見、病態解明につながる研究ができればと思う。

5. 研究総括の見解

行動に対する報酬を予測しこれを最適化する強化学習においては中脳ドーパミン細胞が予測誤差を表現するとされているが、その誤差信号の計算メカニズムの実体は不明であった。本研究では眼球運動(サッケード)課題を学習してジュース報酬予測を行なうサルで多数の中脳ニューロンの活動を記録・解析して、予測誤差計算モデルに必要な3種のニューロンが中脳ドーパミン細胞へ興奮性入力を送る脚橋被蓋核に実際に存在することを示したこと、さらに注



意深い解析により、課題文脈依存性および感覚情報取り込みを円滑にするサッケード抑制と関係するニューロンを発見したことが特に評価される。また背側縫線核が関与する嫌悪・時間割引なども包摂した報酬・嫌悪の強化学習モデルを共同研究により提唱し、実証研究を進めた。一方提案にはなかった研究ではあるが、サッケードの反応時間、注視中の固視微動には、運動の実行、経験に基づく適切な随意運動発現に必要な運動準備状態に関わる情報が含まれてことを発見した。今後はサルを用いた強化学習の神経生理学研究とヒトでの心理物理実験とが統合的に発展し、生物の行動の基本原理の解明につながることを期待される。

6. 主な研究成果リスト

(1) 論文(原著論文)発表

1. 著者. 発表論文タイトル. 掲載誌名. 発行年, 巻号, 始頁-終頁, その他

1. Okada K and Kobayashi Y Characterization of oculomotor and visual activities in the primate pedunclopontine tegmental nucleus during visually guided saccade tasks. *European Journal of Neurosciences* 30: 2211-2223, 2009
2. Kobayashi Y Reward prediction error computation in the pedunclopontine tegmental nucleus neurons. *The 3rd International Symposium on Mobiligence*, Nov 19-21, Awaji 2009 pp 153-156
3. Okada K, Nakamura K, Kobayashi Y. A neural correlate of predicted and actual reward-value information in monkey pedunclopontine tegmental and dorsal raphe nucleus during saccade tasks. *Neural Plasticity* 2011:21, 2011.
4. Okada K, Kobayashi Y. Reward prediction-related increases and decreases in tonic neuronal activity of the pedunclopontine tegmental nucleus. *Frontiers in Integrative Neuroscience*, 14 May 2013 | doi: 10.3389/fnint.2013.00036
5. Masayuki Watanabe, Yuka Matsuo, Ling Zha, Douglas P. Munoz, and Yasushi Kobayashi Fixational saccades reflect volitional action preparation. *Journal of Neurophysiology*, 2013 *in press*

(2) 特許出願 研究期間累積件数:0 件

(3)

その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

国際学会招待講演

Yasushi Kobayashi Neuronal recording from cholinergic nucleus in the brainstem of awake monkeys in relation to arousal, motivation, motor control, visual information processing and reinforcement learning, Bio-X 2012 International Symposium on Molecular Cognition and Translational Research of Neuropsychiatric Disorders, April 28-30, 2012 in Shanghai Jiao Tong University 上海交通大学, China



Yasushi Kobayashi Neural mechanism of reinforcement learning and motivation.
Shanghai Jiao Tong University BioX-project, July 24, 2011 in Shanghai Jiao Tong
University, China

国内学会招待講演

小林康 北海道大学電子科学研究所 陰的制御研究会
運動誤差学習と強化学習、小脳と脳幹の神経振動 2011 6/28-29

Yasushi Kobayashi

Computation of reward signals in the pedunclopontine tegmental nucleus
neurons. Satellite Symposium of the 32nd Annual Meeting of the Japan Neuroscience
Society The Basal Ganglia in Health & Disease (2009)

国際学会口頭発表

K.-I. OKADA, Y. KOBAYASHI

Sustained tonic excitation and suppression of activity on
pedunclopontine tegmental nucleus neurons in behaving monkeys
SfN 2012/10

K.-I. OKADA1, Y. KOBAYASHI

Saccade-related modulation of rhythmic firing pattern on pedunclopontine tegmental
nucleus neurons in behaving monkeys.
699.16/QQ9 Neuroscience2011, Washington DC Nov 15, 2011

Ken-ichi Okada, Yasushi Kobayashi

Context dependent firing regularity of pedunclopontine tegmental nucleus of
behaving monkeys.
2010 Society for Neuroscience Abstract 591.5

Kazuko Hayashi, Kazuko Nakao, Ryuichi Matsuzaki, Ken-ichi Okada, Yasushi Kobayashi
and Kae Nakamura

Neural processing of appetitive and aversive stimuli in the primate raphe nucleus.
2010 Society for Neuroscience Abstract 511.26

K. Okada & Y. Kobayashi



Analyses of the time course of neuronal activity of the pedunculo-pontine tegmental nucleus in monkeys for reward conditioned saccade task
Society for Neuroscience Meeting, 583.17, 2009

K. Nakao, R. Matsuzaki, A. Noritake, K. Okada, Y. Kobayashi and K. Nakamura
Positive and negative value coding in the primate dorsal raphe nucleus
Society for Neuroscience Meeting, 683.13, 2009

Yasushi Kobayashi

Reward prediction error computation in the pedunculo-pontine tegmental nucleus neurons. The 3rd International Symposium on Mobiligence,
Nov 19-21, Awaji 2009

Yasushi Kobayashi and Ken-ichi Okada

The pedunculo-pontine tegmental nucleus neurons relay predicted and actual rewarded information. The international symposium on Construction and Reconstruction of the brain, Oct 8-10, Awaji 2009 pp 55

国内学会口頭発表

平成23年度 生理学研究所研究会 2012/01

グローバルネットワークによる脳情報処理

「サル脚橋被蓋核における周期的活動とその変調」

岡田 研一 小林康

日本神経科学会 2012/09

サル脚橋被蓋核ニューロンにおけるサッカード抑制応答

Saccade-related pause and rebound of activity on pedunculo-pontine tegmental nucleus neurons in behaving monkeys

岡田 研一、小林 康

第15回視覚科学フォーラム研究会

2011年8月29・30日

サル脚橋被蓋核におけるサッカードに伴う周期的活動の変調

岡田研一・小林康

第5回 Motor Control 研究会

2011年6月16～18日

サル脚橋被蓋核ニューロンの周期発火とサッカード抑制応答

岡田 研一、小林 康



Yasushi Kobayashi Ken-ichi Okada The pedunculo pontine tegmental nucleus neurons encode predicted reward signal by tonic regular firing and given reward signal phasically. Neuro 2011 Yokohama 9/16 03-E-1-4

Yasushi Kobayashi, Ken-ichi Okada
The pedunculo pontine tegmental nucleus neurons relay predicted and actual reward and context dependent visuomotor information.
Neuro2010 Meeting Sep 2-4, Kobe 01-8-1-1

Yuri Kitamura, Yuka Matsuo, Masako Taniike, Ikuko Mohri, Yasushi Kobayashi, Kanehisa Morimoto
Saccadic eye movements as a neural correlate measure of preparatory set in children with ADHD.
Neuro2010 Meeting Sep 2-4, Kobe 02-7-3-3

Ken-ichi Okada, Yasushi Kobayashi
Rhythmic firing of pedunculo pontine tegmental nucleus neurons in behaving monkeys.
Neuro2010 Meeting Sep 2-4, Kobe 02-9-4-1

Kazuko Hayashi, Kazuko Nakao, Ken-ichi Okada, Yasushi Kobayashi, Kae Nakamura
Neuronal coding of rewarding and aversive stimuli in the primate dorsal raphe nucleus.
Neuro2010 Meeting Sep 2-4, Kobe P3-113

Yasushi Kobayashi & Ken-ichi Okada
Computational mechanism of reward prediction error by the Pedunculo pontine tegmental nucleus neurons.
The 32nd annual Meeting the Japan Neuroscience Society 01-J1-4 S48 (2009)

Ken-ichi Okada & Yasushi Kobayashi
Relation of pedunculo pontine tegmental nucleus neurons in monkeys to reward prediction and behavior.
The 32nd annual Meeting the Japan Neuroscience Society P1-j08 S114 (2009)

Kazuko Nakao, Ryuichi Matsuzaki, Ken-ichi Okada, Yasushi Kobayashi, Kae Nakamura
Reward coding by the primate dorsal raphe neurons is context dependent
The 32nd annual Meeting the Japan Neuroscience Society P1-j03 S113-114 (2009)

受賞



2010 年 日本神経回路学会論文賞受賞

書籍

小林康 強化学習の神経機構 ブレインサイエンスレビュー2012 クバプロ ISBN:
978-4-87805-121-0

Yasushi Kobayashi and Ken-ichi Okada Book title: *Advances in Reinforcement Learning*
(2010) Open Access Publisher INTECH, ISBN: 978-953-307-369-9 Chapter Title: Reward
prediction error computation in the pedunclopontine tegmental nucleus neurons pp
157-180